REFERENCE ARCHITECTURE

AIRI® for AI Workload Scaling

Get faster time-to-insights with Pure Storage® FlashBlade® and NVIDIA DGX A100.

Contents

Introduction

Executive Summary

AI-Ready Infrastructure (AIRI®) is an end-to-end infrastructure solution from Pure Storage® and NVIDIA for artificial intelligence and deep learning (DL). The NVIDIA DGX A100 system is a quantum leap in AI and machine learning (ML), and it requires a highly performant storage system to ensure GPUs are saturated with data, especially at multi-system scale. Pure Storage FlashBlade® is both performant and scalable for both file and object protocols and maintains its management simplicity across all sizes of deployments.

We evaluated the AIRI reference architecture with Pure Storage FlashBlade and NVIDIA DGX A100 systems for suitability for supporting DL workloads from experiments to production.

Testing at an NVIDIA data center demonstrated near-linear scaling of top-tier performance results across a variety of workloads. This document describes the results from these tests and discusses how AIRI uniquely meets the infrastructure requirements of AI platforms. To learn more about the system architecture of the system and configuration steps, refer to the AIRI POD Configuration Guide.

Introduction

Advances in deep neural networks have ignited a new wave of applications for AI. Powerful new tools and techniques have enabled breakthroughs in fields as diverse as autonomous vehicles, natural language processing, and predictive health care.

Designing and configuring infrastructure to enable large-scale DL requires a significant investment of time and resources to avoid unforeseen delays, bottlenecks, or downtime. AI teams are overloaded with information, and decisions at the infrastructure architecture stage can make or break AI projects.

NVIDIA and Pure Storage have joined forces to deliver a fully integrated platform that offers scale-out DL out of the box, with time-to-insight in hours rather than days or weeks.

About the NVIDIA DGX A100 System

The DGX A100 system is the universal system for all AI workloads, offering unprecedented compute density, performance, and flexibility in the world's first 5 petaflop AI system. Built on the revolutionary NVIDIA A100 Tensor Core GPU, the DGX A100 system unifies data center AI infrastructure, running training, inference, and analytics workloads simultaneously with ease. More than a server, the DGX A100 system is the foundational building block of AI infrastructure and part of the NVIDIA end-to-end data center solution created from over a decade of AI leadership by NVIDIA. The DGX A100 system integrates exclusive access to a global team of AI-fluent experts that offer prescriptive planning, deployment, and optimization expertise to help fast-track AI transformation.

About Pure Storage FlashBlade

Pure Storage developed our FlashBlade architecture to meet the storage needs of data-driven businesses. FlashBlade is an all-flash system, primarily optimized for storing and processing unstructured data. A FlashBlade system can simultaneously host multiple file systems and multi-tenant object stores for thousands of clients. FlashBlade is a scale-out, all-flash storage system, powered by a distributed file system that is purpose-built for massive concurrency across all data types. It can scale up to multi-petabyte capacity with linear-scale performance simply by adding a single blade at a time, up to 150 blades. Because of its native scale-out architecture and ability to drive performance for any type of workload, FlashBlade enables enterprises to consolidate a range of workloads, from backup to analytics and AI, on a single platform.

Storage Requirements for an AI Platform

The key metric in filesystem performance for DL training is read performance. During DL training, the system repeatedly reads data into compute nodes while iterating the model iterated. Pure Storage FlashBlade delivers consistent linear performance as capacity and performance needs for the system scale, allowing for accurate budgeting when forecasting future storage capacity and performance needs as data sets and training algorithms increase in size and sophistication.

However, taking DL models to production involves more than just DL training workloads. An AI platform should be performant across a wide range of workloads such as data investigation, inference, and general analytics. FlashBlade consistently performs across these diverse workloads, which enables other teams within the organization to utilize the same central data infrastructure, minimizing data duplication and data management. In addition, FlashBlade provides high performance for both file and object workloads.

Existing storage systems sacrifice one or more of these dimensions or force architects and administrators to suffer through excessive deployment and management complexity.

Expert-level administrators can tune parallel file systems designed specifically for the needs of high-performance computing (HPC) to meet workload requirements. However, a new data set or training paradigm inevitably requires a new configuration and tuning process, resulting in project delays and potential stranded capacity.

Traditional NAS offerings can provide strong resilience and scalable capacity but often fail to deliver the performance required across a range of I/O patterns and large-scale compute clusters.

FlashBlade is the only storage solution that meets all the AI workload requirements while keeping configuration and management complexity to a bare minimum. The FlashBlade scale-out architecture is purpose-built for massive concurrency across all data types—perfect for centralizing data for all workloads across an AI platform.

AIRI enables seamless scaling of both NVIDIA DGX A100 systems and Pure Storage FlashBlade and is designed for teams solving for faster time-to-insight.

AIRI FOR MLOps

AI platforms need to do more than just support DL frameworks like TensorFlow and PyTorch; they must support more than just model development workloads. An AI platform needs to provide testing pipelines, versioning, sandbox environments, monitoring, and more.

For example, you might start creating Kubernetes clusters for AI workloads. That cluster will run a wide set of applications that need access to a variety of datasets and compute hardware—and likely even a variety of protocols.

Pure includes all FlashBlade software which means users have access to the Pure Service Orchestrator. Pure Service Orchestrator fully automates the creation and management of PersistentVolumes (PV) for applications in a Kubernetes cluster.

Infrastructure simplicity lowers your threshold for shipping models into production. With the AIRI reference architecture from Pure Storage and NVIDIA, you can now support the end-to-end AI lifecycle from development to deployment on one elastic infrastructure.

Architectural Overview

The AIRI architecture was designed around two primary considerations:

- Building a dynamic and agile infrastructure to support the iterative nature of AI projects
- Future-proofing both compute and storage investments as companies mature their AI-based solutions and capabilities

The tests discussed in this document were run with the following AIRI configuration:

Component	Configuration	Networking
4x NVIDIA DGX A100	32x A100 GPU (total)	8x 200GB/s adapter for GPU-GPU traffic 2x 200GB/s for storage traffic 2x 200GB/s for management traffic
1x Pure Storage FlashBlade	30x 17TB blades (301.7TB usable, before data reduction)	16x 100GB/s uplinks from External Fabric Modules (XFMs)

Table 1: Tested AIRI configuration.

The portfolio of AIRI reference architectures supports top-of-rack networking from enterprise companies like NVIDIA, Arista, and Cisco. For further information about switch configuration in an AIRI deployment, see the AIRI POD Configuration Guide.

The AIRI architecture is designed for scale-out DL workloads and is not restricted to the size of this test configuration. As compute demands grow, additional DGX A100 systems can be provisioned in the high-performance fabric and instantly access all available datasets. Similarly, as storage capacity or performance demands grow, additional blades can be added

to the FlashBlade system with zero downtime or reconfiguration. See our Hyperscale AIRI documentation for networking best practices for scaling AIRI with up to 512 DGX A100 systems.

Test Methodology

This section presents benchmark results from a variety of benchmark tests we ran on the AIRI platform. Our testing answers the questions "How does the overall read throughput of the system scale?" and "How does the system perform during real-world AI training workloads?"

The NVIDIA Collective Communications Library (NCCL) tests verify the maximum scalability of GPU-to-GPU communication across multiple DGX A100 systems. Multi-node training jobs require high-performance RDMA communication between nodes. The results of this test showcase the ability of the complete AIRI POD solution to scale multi-node traffic.

The filesystem testing tool fio is used to test single- and multi-node bandwidth performance. Read performance is critical for DL training. Data loading is often a bottleneck in real-world training jobs, so high, scalable read performance is imperative or else storage can become the bottleneck during training.

MDtest measures the metadata performance of a file system. As more data scientists explore, subset, and create datasets, it is critical for AI infrastructure to scale metadata performance.

Finally, MLPerf is an industry-standard set of benchmarks for neural networks. We ran a sample MLPerf workload to demonstrate performance scalability during a real-world end-to-end training job.

We ran these tests to validate the simplicity of the AIRI architecture and demonstrate that the AIRI system provides near-linear throughput as the tests scale. Across tests, as the number of DGX A100 systems increased, the FlashBlade system provided predictable increases in performance while maintaining management simplicity; after initial installation, the FlashBlade system required no tuning or adjustments.

The following tests were conducted in an NVIDIA data center. The tests verified expected system performance with an AIRI system out of the box and with no infrastructure tuning. For more specifics about architecture and configuration, please refer to the AIRI POD Configuration Guide.

Optional NFS Mount Settings

Traditionally, an NFS client is bottlenecked by the throughput limits that a single TCP connection can provide, i.e., the maximum bandwidth from a single NIC on the host to a single blade on the storage server.

To take advantage of the full potential of the FlashBlade system and utilize more than one blade to serve traffic to a single host, we utilized the nconnect mount functionality so that each DGX A100 system could use multiple TCP connections (up to 16) to FlashBlade. Thus, in this testing environment where the FlashBlade system contains 30 blades, the maximum total read bandwidth across the system is about 30 GB/s.

Here is an example command for using the nconnect setting during mount:

```
mount -t nfs -o ro,nconnect=16 198.18.0.100:/datasets /mnt/datasets
```

NCCL all_reduce_perf Scalability Test

NCCL tests verify scalability across multiple DGX A100 systems. Within a system, the expected network bottleneck is the bandwidth of the NVIDIA NVLink® high-speed interconnect. Across multiple systems, the expected network bottleneck occurs at the RoCE-enabled Ethernet adapters that are assigned for GPU-to-GPU communication across DGX A100 systems.

For a single DGX A100 system, optimal results for this test would be for inter-GPU bandwidth to reach NVLink interconnect capabilities. For multiple DGX A100 system configurations, the optimal result would be inter-GPU bandwidth reaching aggregate bandwidth of all Ethernet adapters assigned to the test.

For a comparable NCCL test suite, please see the NVIDIA nccl-tests GitHub repo.

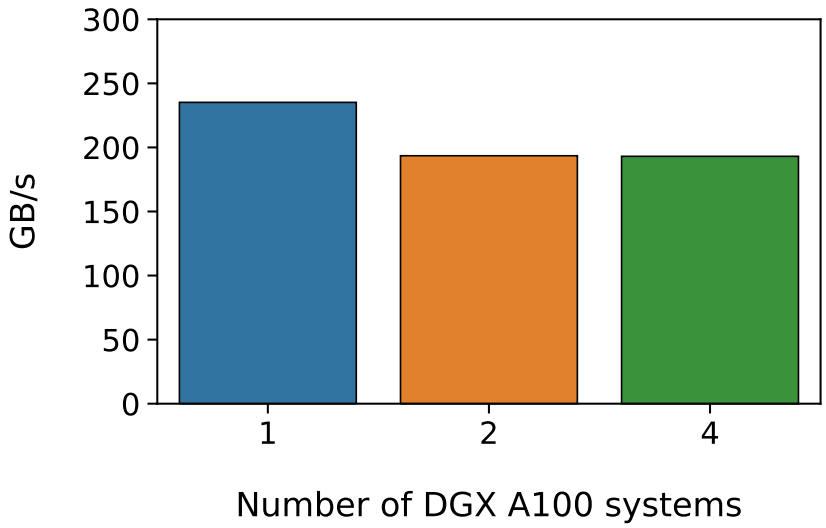

Figure 1: Test results for NCCL all_reduce_perf tests with AIRI. AIRI enables maximum inter-GPU network bandwidth.

Fio Bandwidth Test

fio is a well-known storage microbenchmarking tool. It provides reliable and reproducible storage I/O for local storage and external or distributed storage targets. It also supports the ability to scale testing across multiple clients.

The optimal results for this test would be a linear scaling of FlashBlade bandwidth as the number of DGX A100 systems increases.

With a sample synthetic workload produced by fio, four DGX A100 systems were easily able to drive over 25 GB/s read traffic from the FlashBlade system.

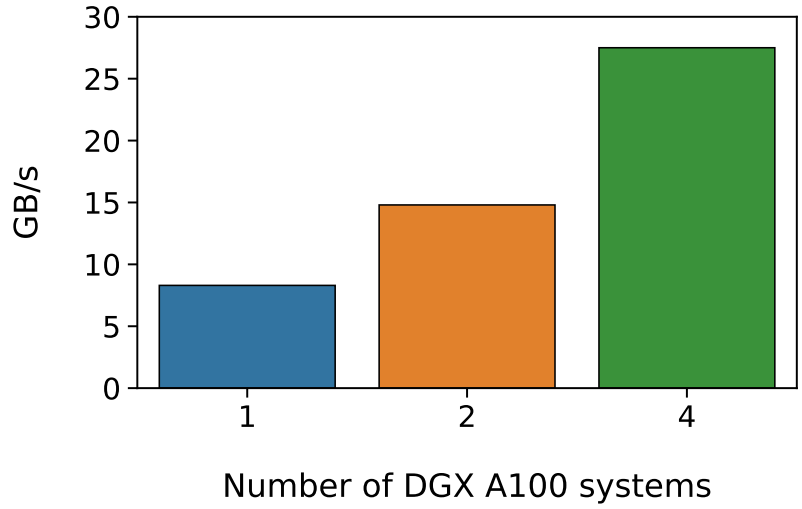

Figure 2: Test results for fio with AIRI demonstrating scalability read bandwidth.

FlashBlade can seamlessly provide high throughput to each DGX A100 system with no tuning. The 6+ GB/s system read throughput per node demonstrated in this testing exceeds the storage performance most organizations will need for their read bandwidth during DL training.

MDTest Metadata Test

While DL training is a heavy read workload, data science teams using an AI platform also need to perform tasks that generate other I/O patterns on the infrastructure. For example, data exploration tasks and dataset subsetting often drive high metadata ops with the "ls" or "find" command. With AIRI, developers have fast system performance no matter the task. As an example, we demonstrate the scalability of the metadata ops file creation and file stat.

MDtest measures the metadata performance of a file system. In this test, we used the program to create a directory tree as well as to time file creation and file stat metadata ops.

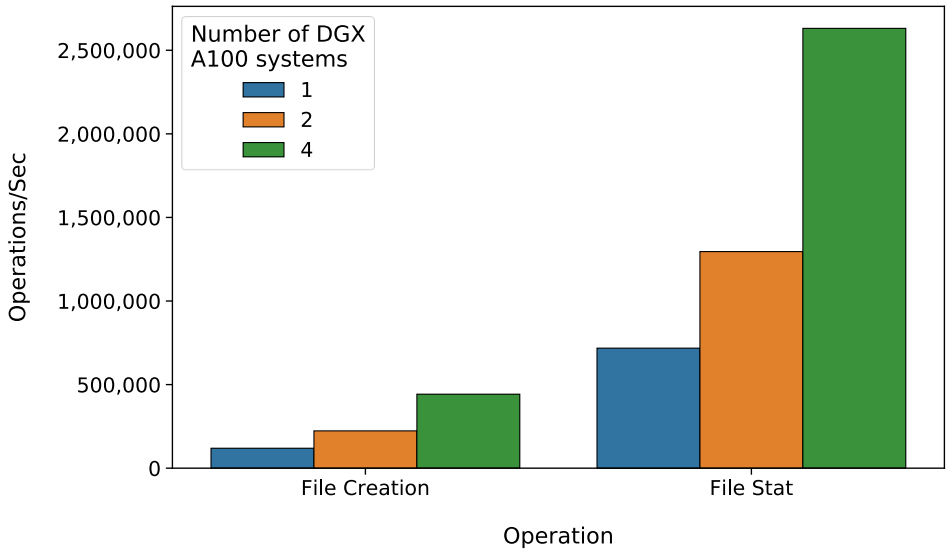

Figure 3: Test results for MDTest with AIRI demonstrating the scalability of file creation and file stat metadata ops.

AI teams want to be able to anticipate their system's scaling. DGX A100 systems and FlashBlade make that possible by giving clear and predictable scaling points. Whether it is multiple large training jobs or data science as a service or both, *AIRI provides predictable performance and simplicity.*

MLPerf and ResNet-50 Performance

MLPerf is an industry-standard set of benchmarks for neural networks. We tested the AIRI system with the ResNet-50 neural network as it is a well-known image classification network that can be used with the ImageNet dataset. It is sufficiently computationally-intensive and is the most I/O-intensive workload.

Methodology

NVIDIA built the MLPerf v0.7 implementation of ResNet-50 with the MXNet framework. The training data is ImageNet, formatted using RecordIO. The results presented have a consistent batch size per system of 408 images as the workload is scaled (weak scaling).

MLPerf tests were run with one, two, and four DGX A100 systems. Slurm, NVIDIA's Pyxis plugin, and Enroot software were used to coordinate the work across all DGX A100 systems involved in the task. Please see NVIDIA DeepOps documentation for more details on how the configuration was set up.

The base container image was the 20.06 MXNet image from NVIDIA NGC. The NVIDIA DALI framework is used to accelerate decoding and ingesting data into the A100 GPUs. Horovod is used to coordinate the training workload across multiple DGX A100 systems.

For steps to launch a training run using the mlperf-nvidia:image_classification docker image, please refer to this documentation. For environment variables specific to this test run, please refer to the AIRI POD Configuration Guide.

Results

Ideally, a training job that is not bottlenecked by storage should be able to scale images/sec throughput linearly as more DGX A100 systems participate in the job.

Figure 4 shows that, as node count increases, images per second throughput for AIRI continues to be highly performant and scale predictably. FlashBlade's performance maintains high-read bandwidth so that storage is not a bottleneck during training.

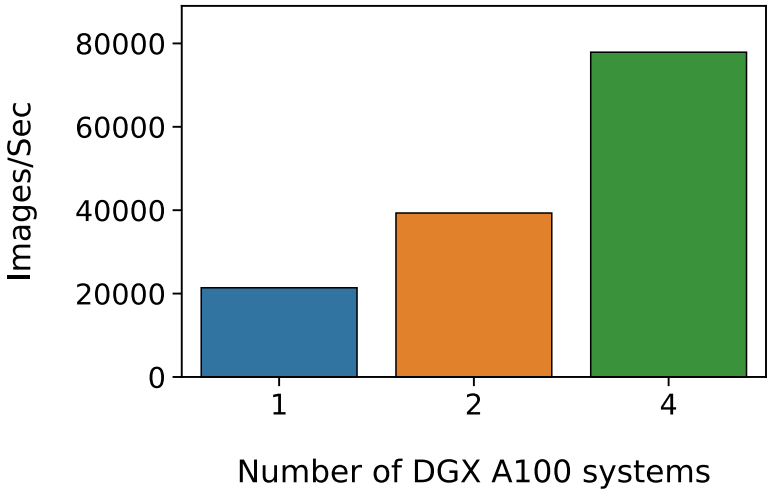

Figure 4: AIRI throughput performance and scalability

Additionally, one of the common pitfalls when evaluating infrastructure for deep learning is to only measure performance at Epoch 0 of a training job. It is critical to also evaluate the training job's "steady state"–for example, also measuring performance at Epoch 20.

Traditional storage solutions are not able to keep up with the read bandwidth demands of a training job, and the data load bottleneck becomes more apparent as the job continues. If a job's images/sec is lower at Epoch 20 than at Epoch 0, then storage is unfortunately limiting training performance.

As validated during testing with NVIDIA, FlashBlade provides the same high performance at Epoch 20 as at Epoch 0 during a training job's steady state, enabling training jobs to run with maximal throughput for their duration.

*Note: These results reflect the scalability of the end-to-end system, **including the current software stack**. While data loading performance is impacted in part by storage server performance, processing tasks on the compute nodes impacts performance as well. As libraries improve over time, we expect the MLPerf benchmark results for AIRI to evolve as well.*

Conclusion

In the modern era of artificial intelligence, enterprises will tap the power of AI for innovation and competitive advantage. However, many AI initiatives are stalled by infrastructure complexities, often crippled with legacy, do-it-yourself software, and hardware technologies.

AIRI is the industry's complete AI-ready infrastructure, architected by NVIDIA and Pure Storage to extend the power of DGX A100 systems, enabling AI-at-scale for every enterprise.

By using a flexible yet homogenous infrastructure built on one system type, IT teams have the flexibility to adapt to business demands over time. This sets a strong foundation for AI projects with compute and storage that flexibly adapt to changing workload demands.

Fast, simple, and future-proof, AIRI is the ultimate solution to empower data architects and scientists to deliver insight at scale, in hours rather than weeks.

Additional Resources

- Learn how to simplify your AI transformation with AIRI.
- See how to store a private Docker registry on FlashBlade S3.
- Read about scraping FlashBlade metrics using a Prometheus exporter.
- For more details on FlashBlade in an AI environment, visit Pure Storage AI Solutions.

Pure Storage, Inc.
650 Castro Street, #400
Mountain View, CA 94041